



Teste Cientista de Dados CGU

Objetivo

Este teste prático tem como objetivo avaliar suas habilidades técnicas nos requisitos da vaga, bem como compreender seu processo de tomada de decisões técnicas e análise de trade-offs.

Descrição do Desafio

Você deve construir uma API utilizando FastAPI I com três endpoints principais, conforme especificado abaixo.

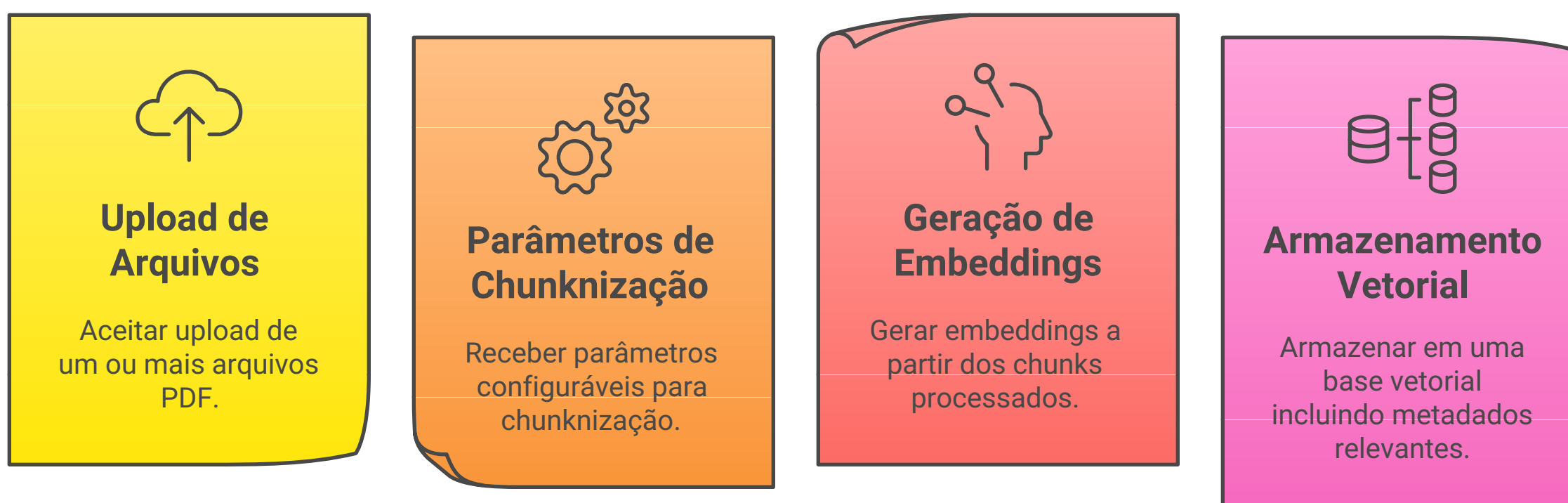
Especificações dos Endpoints

1. Processamento de Documentos

Endpoint: Upload → Chunknização → Embedding → Base Vetorial

Funcionalidades:

- Aceitar upload de um ou mais arquivos PDF
- Receber parâmetros configuráveis para chunknização
- Gerar embeddings a partir dos chunks processados
- Armazenar em uma base vetorial incluindo metadados relevantes



2. naive-RAG

Funcionalidades:

- Receber uma pergunta como input
- Retornar uma resposta adequada baseada nos documentos indexados
- Permitir ao usuário escolher se deseja aplicar reranking com BM25

3. Classificação de Texto

Endpoint: Análise de Sentimentos ou Classificador

Funcionalidades:

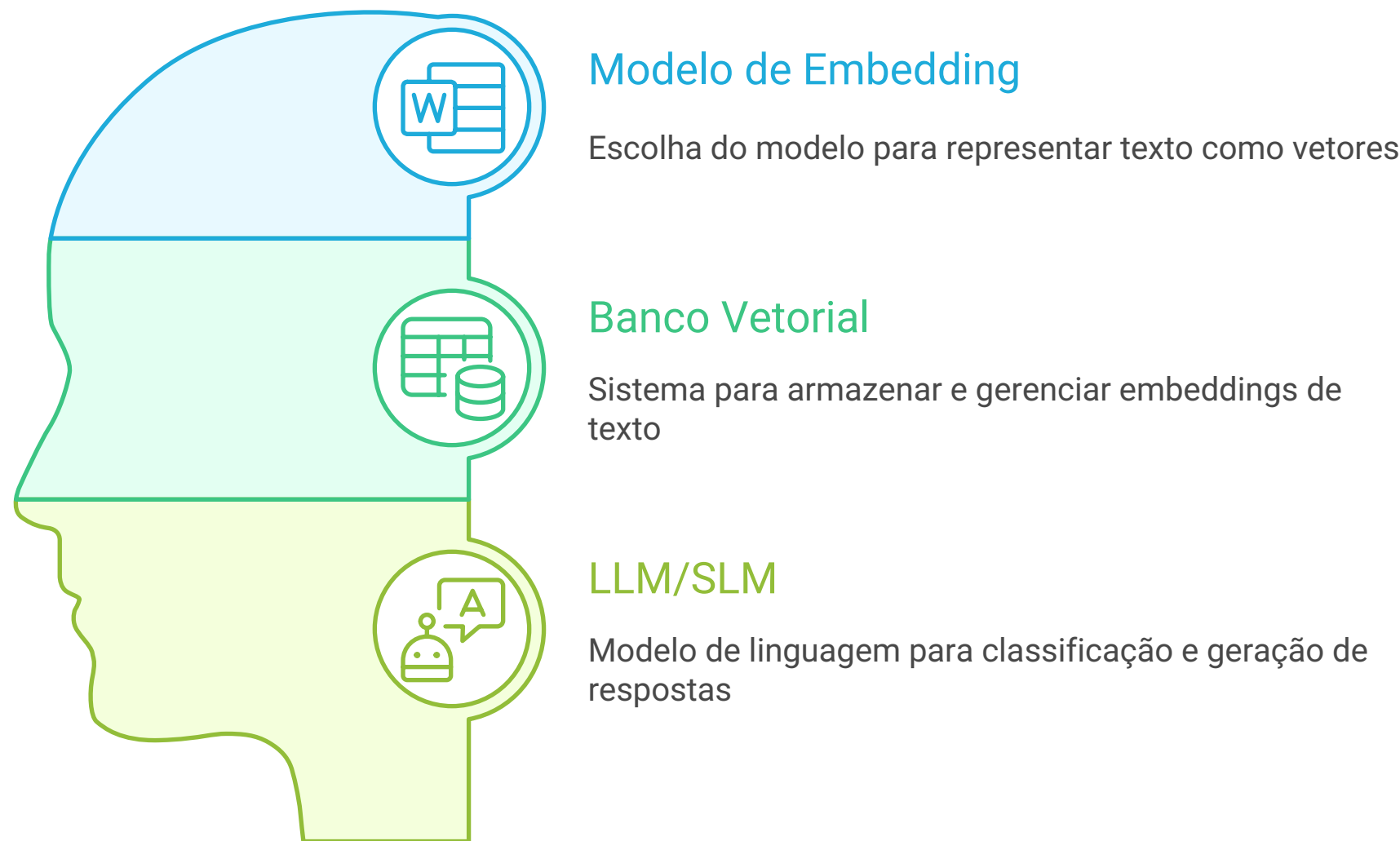
- Receber uma sentença como input
- Classificar adequadamente utilizando LLM ou SLM
- Preferencialmente, utilizar logprobs para fundamentar a classificação

Decisões Técnicas

As seguintes escolhas são parte integrante da avaliação:

- · Modelo de embedding a ser utilizado
- · Banco vetorial para armazenamento
- · LLM/SLM para classificação e geração de respostas

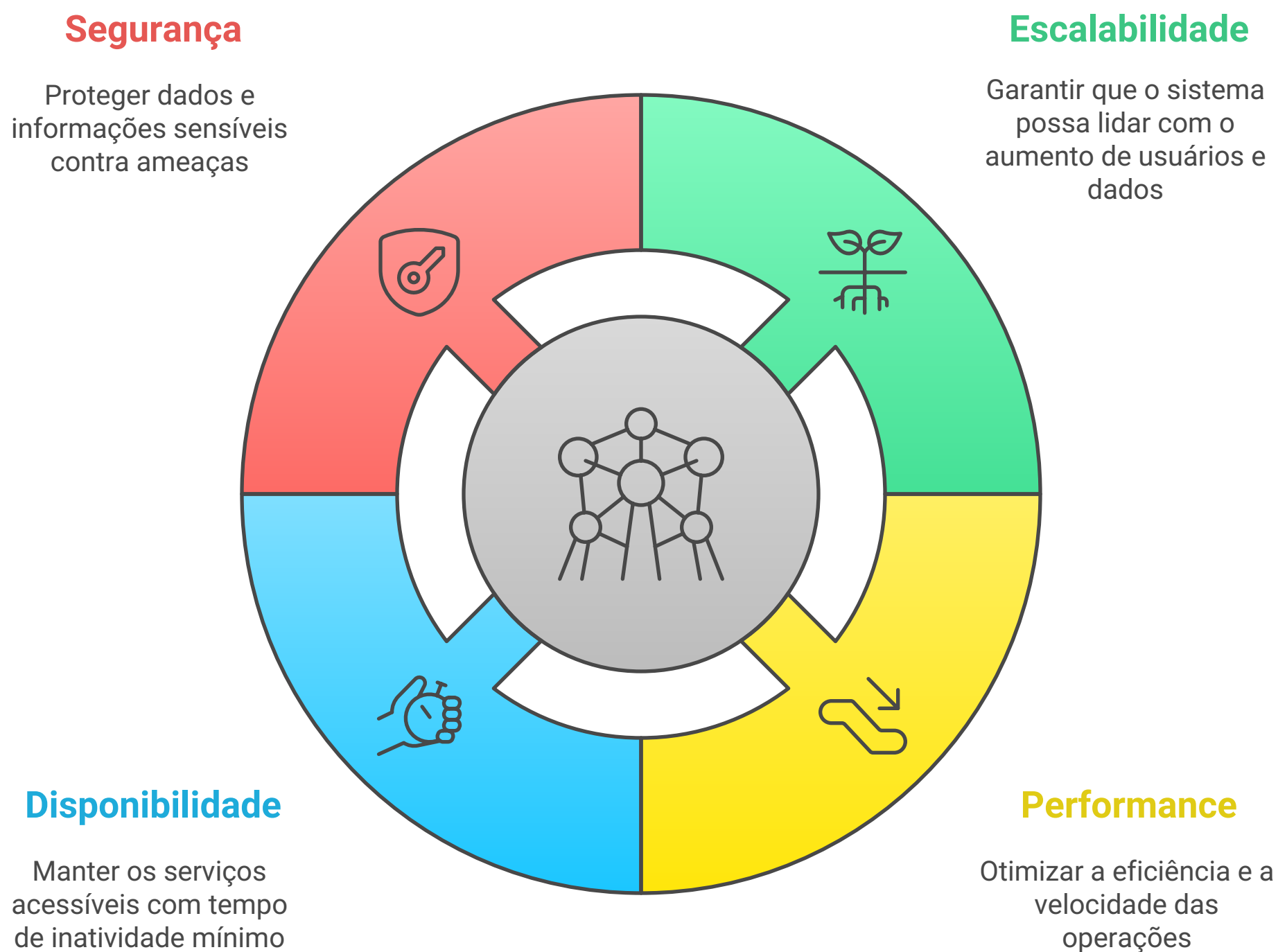
Mas, podem ter outras.



Requisitos de Arquitetura

A API deve ser projetada para suportar até 10.000 usuários concorrentes. Elabore um desenho detalhado da arquitetura considerando:

- · Escalabilidade
- · Performance
- · Disponibilidade
- · Segurança



Entregáveis Obrigatórios

1. [README.md](#) completo e bem documentado
2. Desenho da arquitetura [diagrama + explicação]
3. Repositório Git com o código da solução

Apresentação da solução

A apresentação, de no máximo 15 minutos, deve demonstrar seu raciocínio analítico, destacando as escolhas técnicas realizadas, os desafios enfrentados e como você avaliou diferentes opções para chegar à solução apresentada.

Observações

- Não é necessário entregar o código completo funcional, pode ser um esboço das APIs e serviços
- Não tem restrição quanto ao modelo de LLM utilizado
- Os documentos a analisar podem ser de qualquer tipo, contanto que não infrinjam os direitos autorais
- Foque na qualidade da documentação e no design da solução
- Justifique suas escolhas técnicas no README